AD-A091 679 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF PSYCHOLOGY F/G 5/10
THE EFFECTS OF CATEGORY GENERALIZATIONS AND INSTANCE SIMILARITY--ETC(U)
OCT 80 R ELIO, J R ANDERSON N00014-78-C-0725
UNCLASSIFIED TR-80-6 NL

1 OF 1
AD A
091679

END
DATE
FILMED
12-80
DTIC

LEVEL 12

The Effects of Category Generalizations and Instance Similarity on Schema Abstraction

Renee Elio and John R. Anderson

Department of Psychology
Carnegie-Mellon University
Pittsburgh, Pennsylvania

This research was supported in part by an NSF graduate fellowship to the first author and by ONR contract No. N00014-78-C-0725 to the second author. The authors wish to thank Gary Bradshaw, Bill Jones, Lynne Reder, and Miriam Schustack for their comments and suggestions on all phases of this study, Paul Kline for helpful criticism of an earlier version of this manuscript, and Mark Premo for running subjects for Experiment III. Requests for reprints should be sent to Renee Elio, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

unclassified (14) TR-

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report #80-6	2. GOVT ACCESSION NO. AD-A091679	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Effects of Category Generalizations and Instance Similarity on Schema Abstraction		5. TYPE OF REPORT & PERIOD COVERED (9) Technical ~~report~~,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Renee Elio ~~and~~ John R. Anderson		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-78-C-0725
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Carnegie-Mellon University Pittsburgh, Pennsylvania 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (17) RR042-04-01 NR 154-399 66153N
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE (11) 7 Oct~~ober 7~~, 1980
		13. NUMBER OF PAGES 53 (12) 53
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (16) RR04204		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Prototypes	Instance similarity
Learning	Category formation
Generalization	Spacing
Production systems	Partial matching
Schema abstraction	Analogy

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Three experiments were designed to differentiate two models of schema abstraction. One model, called the generalization model, proposes that category generalizations, defined as feature combinations which occur frequently across study items, are abstracted during learning and used to classify transfer items. According to the other model, called the instance-only model, transfer items are classified according to their similarity to studied items. Study materials were constructed which either yielded category generalizations (generalize condition) or did not (control condition). Transfer items differed

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

301 ...

unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. Abstract (Continued)

on whether they were classifiable by category generalizations and on their similarity to study items. In Experiments I and III, accuracy and confidence on transfer items was better in the generalize condition than in the control condition. Experiment II manipulated the order in which generalizable study items were presented for study: Items were either blocked, so that items contributing to a category generalization occurred close in the study sequence, or randomly ordered. Study items were learned faster and transfer performance was better with blocked presentation than with random presentation. In all three experiments, there was an effect for the similarity of transfer items to study material. There was some evidence suggesting an advantage for partially matching a category generalization. The results support a schema abstraction model in which transfer is a function of similarity to both specific category instances and to higher-order category information.

unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Abstract

Three experiments were designed to differentiate two models of schema abstraction. One model, called the generalization model, proposes that category generalizations, defined as feature combinations which occur frequently across study items, are abstracted during learning and used to classify transfer items. According to the other model, called the instance-only model, transfer items are classified according to their similarity to studied items. Study materials were constructed which either yielded category generalizations (generalize condition) or did not (control condition). Transfer items differed on whether they were classifiable by category generalizations and on their similarity to study items. In Experiments I and III, accuracy and confidence on transfer items was better in the generalize condition than in the control condition. Experiment II manipulated the order in which generalizable study items were presented for study: Items were either blocked, so that items contributing to a category generalization occurred close in the study sequence, or randomly ordered. Study items were learned faster and transfer performance was better with blocked presentation than with random presentation. In all three experiments, there was an effect for the similarity of transfer items to study material. There was some evidence suggesting an advantage for partially matching a category generalization. The results support a schema abstraction model in which transfer is a function of similarity to both specific category instances and to higher-order category information.

The Effects of Category Generalizations and Instance Similarity on Schema Abstraction.

It is a ubiquitous phenomenon that people are able to detect regularities that characterize a category of stimuli simply from experience with category members. The success of this inductive process is not limited to well-defined categories, those for which a single rule or list of defining attributes will always predict category membership. For most of the real-world categories, there may be several complex rules governing membership, none of which is singularly predictive. We will call the process by which people learn ill-defined categories from experience with exemplars schema abstraction. We differentiate this process from concept identification only because this latter term has traditionally denoted classification learning situations in which the categories are defined by a single rule, often derived through explicit hypothesis testing. Since the acquired information abstracted from ill-defined categories does not reduce to a simple, easily specified rule, the general issue which concerns us is the nature of that information and how it is subsequently used to differentiate category members from non-members. In the usual schema abstraction paradigm, subjects first learn to classify a set of training items into one or more categories by trial and error. They are then given a set of transfer items, items which they had not studied during training, to assign to one of the categories they learned, usually without feedback. It is their performance on these transfer items which allows us to infer something about the nature of the category information acquired from experience with the initial set of training items.

Models of schema abstraction differ primarily in their conception of the nature of this information, its representation, and its utilization to classify new exemplars. According to prototype models (Posner & Keele, 1968; Franks & Bransford, 1971), a single representation of the category's central tendency, called a prototype, is abstracted during learning as the average of the seen exemplars. Instances are categorized according to how close they are to the prototype. This model accounts for the abstraction phenomena that (a) never-studied category prototypes are more likely to be recognized and correctly classified than other, never-studied items (Posner & Keele, 1968), (b) after delay, never-studied category prototypes are better classified than much-studied training exemplars (Posner & Keele, 1970), and (c) classification and recognition of new items is a function of the distance that the item is from its category prototype (Bransford & Franks, 1971; Franks & Bransford,

1971). These findings suggest two characteristics of the information abstracted from category exemplars. First, it is abstracted during experience with the exemplars, rather than computed at test time, since it is available after delay, while specific instance information is not. Second, the availability of this information after delay also suggests it is qualitatively different than instance information.

An alternative theory of schema abstraction is the view that the abstracted category information is based on the frequency with which features and feature combinations occur across exemplars of a category (Reitman & Bower, 1973; Neumann, 1974; Hayes-Roth & Hayes-Roth, 1977). We will refer to these as *strength* or *frequency* models. Hayes-Roth and Hayes-Roth (1977) proposed that the frequency of occurrence of all an exemplar's single features plus all possible combinations of these features (called property sets) comprise the exemplar's representation. The frequency with which a property set occurred among all the encoded exemplars of a category determines its associative strength to that category. They propose that recognition of an exemplar is governed by the associative strengths of its property sets to the category or categories studied. The diagnosticity of a *property set* for a given category was defined as an increasing function of its associative strength to that category and a decreasing function of associative strength to the alternative categories. Stimulus sets can be created where instances which are farther from the central tendency or prototype have higher property-set diagnosticity than instances which are closer to the prototype. Hayes-Roth and Hayes-Roth demonstrated with such material that property set diagnosticity, not prototypicality, predicted classification behavior.

As different as these models may seem, they share the assumption that some information, qualitatively different from the representation of individual instances, is abstracted, stored, and utilized in subsequent recognition and classification judgments. In contrast, Medin and Schaffer (1978) argued that a model positing only one level of information, instance information, can account for the previous schema abstraction results. In a series of experiments, they controlled the distance of transfer items to the prototypes of two categories while manipulating the similarity of the transfer items to individual category members. Medin and Schaffer demonstrated that the inter-item similarity of training exemplars affected learning time and that subsequent recognition and classification ratings of new instances were a function of their similarity to individual training exemplars, not of their

distance from category prototypes. They propose that the item most similar to a new instance is retrieved and its category assignment is used to classify the new instance.

Medin and Schaffer noted that the method of generating stimuli in most classification learning experiments--creating category exemplars by applying distortion or transformation rules to the category prototype--causes the prototype to be the transfer item most similar to members of its own category and least similar to members of another category. Thus, the similarity-to-stored instances model can account for superior performance on prototypes and items close to prototypes without positing an additional, qualitatively different level of information. It can also account for the result that prototype classification suffers little with delay, because even if some specific instances are forgotten, other instances similar to the prototype will remain.

However, strength models of the feature-set variety can account for the data offered as evidence for similarity-to-stored-instances models because they propose that individual instances are augmented with, not replaced by, higher-order category information. While Medin and Schaffer effectively demonstrated the inadequacy of a prototype model of abstracted category information, their experiments were not designed to contrast the assumptions of their similarity-to-stored-instances model with the assumptions of strength models. The purpose of the present series of experiments is to distinguish between an instance-only model which proposes that transfer performance is a function of similarity to stored exemplars and a particular strength model proposing that some higher-level, qualitatively different information is abstracted from and represented in addition to specific instances, and utilized to classify new items.

The model we are contrasting with the instance-only model is called the *ACT generalization model*, based on Anderson's ACT theory (Anderson, 1976; Anderson, Kline, & Beasley, 1979), a model and a computer simulation of declarative and procedural knowledge. Using general learning mechanisms and assumptions not designed specifically for schema abstraction tasks, the ACT program successfully replicated the recognition and classification results of Franks and Bransford (1971), Neumann (1974), Hayes-Roth and Hayes-Roth (1977) and Medin and Schaffer (1978), given their respective tasks and stimuli (Anderson et al., 1979). However, having the generalization model account for these results simply contributes another competing model to the already large set of alternative schema abstraction theories. We designed the present experiments not simply to marshal

support for the ACT generalization model, but to differentiate the predictions of an instance-only model and frequency-based strength models, of which the ACT generalization model is one version.

The ACT Generalization Model

A generalization is a pattern of frequently co-occurring features in a set of data. While less specific than any pattern seen, a generalization captures the regularities across specific items. For example, we might learn that one member of Club 1 is single, Catholic, plays tennis, and works for the government. We might subsequently learn that a second Club 1 member is single, Protestant, plays tennis, and works for the government. While we would store both these specific feature patterns, the generalization we would form that accommodates both these specific descriptions of Club 1 members would be *Club 1 members are single, play tennis, and work for the government.* Since religion differed in the specific descriptions, it is a variable feature in the generalization about Club 1 members. Note that, in addition to the original study instances, the model proposes that only one feature set--the above generalization--is stored. A generalization must contain more constant features than variable features. With this restriction, a 2-feature generalization such as *Club 1 members are single and play tennis*, while common to both descriptions, would not be formed. This contrasts sharply with the Hayes-Roth and Hayes-Roth (1977) model, which predicts all feature subsets will be stored. It should be obvious that, with only a moderate set of 4- or 5-featured stimuli, the number of possible feature combinations would be enormous. The generalization model's restrictive definition of a generalization acknowledges the limited capacity of short term memory and attempts to make the most efficient use of it.

Each time a generalization successfully classifies a specific feature pattern, its representation in memory becomes stronger, e.g., learning that another Club 1 member is Jewish, single, plays tennis and works for the government reinforces the *Club 1 members are single, play tennis, work for the government* generalization. According to the model, each time a pattern of features successfully classifies an item, not only is it strengthened, but any pattern more general but still consistent with it is also strengthened. For example, the first description *single, Catholic, plays tennis, works for the government* could be classified on some later learning trial by matching the specific feature pattern, (*One Club 1 member is single, Catholic, plays tennis, works for the government*), previously stored for this item. This specific pattern would be strengthed. In addition, the generalization consistent with

this pattern, (*Club 1 members are single, play tennis, and work for the government*). while not the basis for this particular classification, would also be strengthened. Over time, then, a generalization will accrue more strength than any of the specific patterns which generated it. In this manner, particularly useful generalizations, those which are consistent with a large number of specific instances, become stronger than generalizations which are consistent with only a few instances. The effect of this successful application of generalizations is that they have a higher degree of memorial strength than specific patterns encountered. This greater strength is reflected in the higher probability that a generalization rather than a specific instance will be accessed to classify instances. In other words, the above Club 1 member descriptions would eventually be categorized by matching the generalization about Club 1 members rather than by matching the specific patterns initially stored for them. Anderson et al. (1979) offer a more detailed description of the mechanisms we have outlined here.

The various schema abstraction results are easily accommodated by the generalization model. The more distant an item is from its category prototype, the less similar it is to the majority of items and the less likely it will be classifiable by generalizations formed from more prototypic items. The generalization model can also account for the facilitative effect of high inter-item similarity among training exemplars (Medin & Schaffer, 1978): If training items from different categories share a high degree of overlap, generalizations between them will not only compete for application, but their strength will be decremented if they miscategorize items during training. The model can also accommodate the Rosch and Mervis's (1975) findings that an item's classification and typicality ratings depend on its *family resemblance*, the degree to which it is similar to items within its category and dissimilar to items in alternative categories. The generalization model predicts that classification performance on transfer items equally similar to study items from alternative categories would be poor, again because generalizations from different categories would be equally likely to match such items.

Experiment I

Our general plan for distinguishing the generalization model and an instance-only model was to manipulate the likelihood of forming category generalizations in two different sets of study exemplars while holding the similarity of transfer items to the two study sets as constant as possible. In this way,

any advantage for having studied the items which yielded generalizations would be attributed, not to a higher degree of inter-item similarity between those items and the transfer set, but rather to the availability of generalizations.

We also manipulated the type of transfer item. One type of transfer item could be classified by applying category generalizations if generalizations had in fact been formed from experience with generalizable study exemplars. The other type of transfer item was not classifiable by category generalizations. According to the ACT theory as developed in Anderson et al. (1978), a category generalization formed during study must completely match a transfer item in order to classify it. Given this "full match" view, the generalization *Club 1 members are single, play tennis, and work for the government* fully matches and would assign to Club 1 a transfer item such as *single, Baptist, plays tennis, works for the government*, but not a transfer item such as *married, Baptist, plays tennis, and works for the government*. For those transfer items which are matched by generalizations, the model assumes that both specific instance information as well as category generalizations information would be available to classify the items. Performance should be better on these items than on transfer items which do not match category generalizations.

The manner in which this general design was realized in Experiment I can be illustrated best with a small portion of the experimental materials. Subjects read 5-feature descriptions of people who belonged to either the "Dolphin Club" or the "Koala Club." Subjects in the *generalize* condition studied descriptions such as

(1) One member of the Dolphin Club is a Baptist, plays golf, works for the government, is college educated, and is single.

(2) One member of the Dolphin Club is a Baptist, plays golf, works for a private firm, is college educated, and is married.

From these exemplars, we anticipated that they would form the generalization that a member of the Dolphin Club is a Baptist who plays golf and is college educated, since these are the features that these two club members have in common. After learning to classify items like (1) and (2) into the Dolphin Club (and other items into the Koala Club), subjects moved to a transfer task in which they were presented with new items like

(3) This person is a Baptist who plays golf, is unemployed, is college educated, and is divorced.

(4) This person is a Baptist who plays tennis, is unemployed, is college educated, and is divorced.

Description (3) is an instance of what we called a 3-overlap transfer item. It overlaps with both of the two study items (1) and (2) on three features and, moreover, on the three features which form the generalization (*Baptist, golf, college*). Therefore, we would expect transfer performance on item (3) to be quite high, since the generalization formed from (1) and (2) matches it completely. In contrast, description (4) only overlaps with the original study items on two features *Baptist, college*. While both of these features are part of the generalization, we would expect a lower probability of classifying this item as a Dolphin Club member, since a generalization must match an item perfectly to be used. Therefore, a 2-feature overlap with a generalization that requires three features should not help.

The other study condition was called the *control* condition. Rather than studying a pair of items like (1) and (2), subjects might study

(5) One member of the Dolphin Club is a Baptist who plays golf, works for a private firm, is high school educated, and is divorced.

(6) One member of the Dolphin Club is a Baptist who plays golf, works for the government, is college educated, and is married.

Note that these study pairs only overlap on two features, *Baptist* and *golf*. According to the exact model put forth in Anderson et al. (1979), this 2-feature pattern would not emerge as a category generalization because it involves less than 50% of the original features. After learning to classify these items, subjects in the control condition were asked to judge the same transfer items as subjects in the generalize condition. Note that transfer item (3) is still a 3-overlap item for subjects who have studied items (5) and (6). It overlaps with (5) and (6) on three features, but a different set of three features for each study item (with (5) on *Baptist, golf, divorced*, and with (6) on *Baptist, golf, college*). According to the view that inter-item similarity governs classification judgment, performance on (3) should not differ depending on whether subjects studied (1) and (2) or (5) and (6), since the overlap of transfer item with study items is the same. However, the generalization point of view predicts an advantage for having studied items (1) and (2), which offered a 3-feature generalization for classifying (3), whereas (5) and (6) do not. Subjects studying items such as (5) and (6) would have to rely on their memory for these specific instances to classify the transfer item (3). To reiterate, transfer item (3) is equally similar to the generalize study items (1) and (2) and to the control study items (5) and (6), but

in the generalize condition, it also matches a category generalization. Note also that transfer item (4) overlaps with (5) and (6) on two features (but a different set of two features for each). Thus (4) is a 2-overlap item in the control condition as well as in the generalize condition.

The other factor manipulated during learning was whether subjects saw pairs like (1) and (2) or like (5) and (6) close together in the study sequence of study items or randomly spread apart. This was the blocked versus random presentation manipulation. We expected the transfer performance of generalization subjects to be better in the blocked condition than in the random condition: If generalizable pairs are close together, they are more likely to be simultaneously available in a working memory for patterns. Generalizable patterns must be in this working memory for the generalization to be formed. In contrast, we did not predict any particular difference between blocked and random conditions with control study materials.

To summarize, the ACT generalization model predicted performance to be best in the generalize-blocked condition on 3-overlap transfer items, since these items are classifiable by category generalizations, and equally poor on all other types of transfer items. An instance-only model predicts no effect of generalize versus control study material, nor does it predict an advantage for blocking. While it would predict an advantage for 3-overlap versus 2-overlap transfer items, it does not predict that this effect would vary with study material (generalize versus control) or blocking. In contrast, the generalization model predicts an interaction of study material with transfer item type, with the largest effect of 3-overlap versus 2-overlap transfer items for subjects in the generalize-blocked condition.

Method

Subjects. Eighty members of the Carnegie-Mellon University community served as subjects. They received Psychology course credit and/or $3.00 an hour for their participation. Twenty subjects were used in each condition: generalize-blocked, generalize-random, control-blocked, control-random. Subjects were randomly assigned to one of the four experimental conditions upon their arrival for the experimental session, which lasted approximately two hours.

Materials and Design. The stimuli were 5-feature descriptions of people to be classified members of one of two clubs. Each feature had four possible values. The five features and their values were: *job*: (1) unemployed, (2) self-employed, (3) government, (4) private-firm; *marital status:* (1) single, (2) married, (3) divorced, (4) widowed; *religion:* (1) Catholic, (2) Jewish, (3) Episcopalian, (4) Baptist;

hobby: (1) tennis, (2) golf, (3) chess, (4) bowling; *education*: (1) grammar school, (2) high school, (3) college, (4) trade school. Each stimulus item can be described symbolically as five digits, one for each feature, with each digit ranging from 1-4 to indicate the specific value of each feature. Given the above assignment of digits to values, for example, the item 43211 could correspond to the description *private-firm, divorced, Jewish, tennis, grammar school.* The design of the study and transfer material, given in Table 1, was specified in terms of these abstract numbers rather than the specific feature values. Independently for each subject, the values of each feature were randomly re-assigned to the digits 1 through 4. The order of the features in the description was also randomly determined for each subject. This means that for one subject, the 11114 item from Table 1 might have been instantiated as *government, single, Baptist, high school, and chess* while for another subject it might have been instantiated as *Baptist, golf, private-firm, college, and married.* Thus, each subject had his or her own randomly generated set of materials.

Table 1 schematically illustrates the stimuli and design of the experiment. A 2 (study set) X 5 (test item type) X 2 (presentation order) design was used. Study set and presentation were varied between subjects. Table 1 shows items from the two study sets, the generalize set and the control set. Pairs of items in the generalize set yielded 3-feature generalizations. The four Club 1 generalizations were *11-1-, 1--22, 4--11*, and *22-4-*. The pairs of study items in the control condition shared only two features in common. For both study sets, there was no value on any feature which could perfectly predict club membership. The values 1 and 2 were diagnostic of Club 1, since they occurred more frequently on each feature than the values 3 and 4. Club 2 items were constructed by interchanging 1's and 4's with each other and 2's and 3's with each other, so that the values 3 and 4 were diagnostic of Club 2. The third feature was irrelevant witn respect to club membership.

Since the critical aspects of the design rest on the relationships between the study item sets and the test items, it is worthwhile to work through an example group of items. The two study sets were constructed so that pairs of items in each set were equated for the amount of overlap they had with pairs of transfer items. The four types of transfer items are defined with respect to a study pair. Transfer items can be classified according to two characteristics of their relationship to their corresponding study pair: amount of overlap (either three features or two features) and the number of diagnostic values (either three values or two values). Overlap was the number of features for which

Table 1

Experiment I:

Generalize and Control Condition

Study Items and Transfer Items

Study Items

Generalize		Control	
Club 1	Club 2	Club 1	Club 2
11114	44441	11112	44443
11212	44343	11223	44332
12122	43433	13122	42433
13222	42333	24222	31333
42311	13244	21311	34244
44411	11144	42411	13144
22343	33212	22444	33111
22441	33114	42342	13213

Transfer Items

3-1(3-overlap)		3-1(2-overlap)		2-2(2-overlap)		2-2(3-overlap)	
Club 1	Club 2	Club 1	Club 2	Club 1	Club 2	Club 1	Club 2
11313	44242	12413	43142	14313	41242	14113	41442
11413	44142	12313	43242	14413	41142	14213	41342
14322	41233	41422	14133	14323	41232	14123	41432
14422	41133	41322	14233	14423	41132	14223	41332
41111	14444	13111	42444	41131	14424	41331	14224
41211	14344	13211	42344	41231	14324	41431	14124
22142	33413	12141	43414	23142	32343	23442	32112
22242	33313	12241	43314	24242	31313	24342	31213

two items had the same values. For example, the item 14322 has a three-feature overlap (3-overlap) with the item 12122 on the first, fourth, and fifth features, since the values 1, 2, and 2, respectively, are the same for both items. The ratio of diagnostic to undiagnostic features for these Club 1 items is three to one (3-1), since the diagnostic values 1 or 2 occur on the first, fourth, and fifth features and the undiagnostic value 4 occurs on the second feature. (Since the third feature was irrelevant with respect to club membership, its value does not enter into the diagnosticity ratio.) The notation used to describe the transfer items first specifies the ratio of relevant diagnostic values to undiagnostic values and then the amount of overlap. The diagnosticity and overlap factors were manipulated orthogonally, so that there were four types of transfer items: 3-1 (3 overlap), 3-1 (2 overlap), 2-2 (2 overlap), and 2-1 (3 overlap).

To help explain the various transfer conditions, let us go through an example of each transfer item type in Table 1. Consider the first Club 1 generalize study pair: 11114 and 11212. These two items yield a 3-feature generalization *11-1-*. The two corresponding 3-1(3 overlap) transfer items (from the first row of transfer items in Table 1) for this study pair are 11313 and 11413. Each of these transfer items overlaps the generalize study pair on three features which also constitute the *11-1-* generalization (*11*3*1*3 and *11*4*1*3). The three diagnostic values in these items are the 1's on the first, second, and fourth features and the undiagnostic value is 3 on the fifth feature. Of the four transfer item types, only the 3-1(3 overlap) transfer items were classifiable by category generalizations. In the 3-1 (2 overlap) transfer items, (12413 and 12313), the ratio of diagnostic to undiagnostic values is three to one (diagnostic values of 1 or 2 on the first, second, and fourth features, and the undiagnostic value of 3 on the fifth feature). However, the *11-1-* generalization does not completely match these items: They overlap on only two features (the first and the forth) with the generalize study pair 11114 and 11212. In the 2-2 (2 overlap) transfer items, 11413 and 14313, there are two diagnostic values (on the second and fifth features). Each of these items has a 2-overlap with the generalize study pair on the first and fourth features. For these items, category membership was equivocal on the basis of either overlap or the number of diagnostic features. Since they were created by changing one of the diagnostic features in the 3-1(2 overlap) items to be undiagnostic, they were assigned to the same category to which these items belonged. The last type of transfer item, 2-2 (3 overlap), was included to balance the manipulation of diagnosticity and overlap. For these items, however, we were able to

create an 3-overlap with only one of the corresponding study items. Thus, the 2-2 (3 overlap) transfer item 14113 has a 3-overlap with the generalize study item 11114, but not with the other study item 11212. The other 2-2(3 overlap) transfer item from this pair, 14213, has a 3-overlap with 11212 but not with 11114.

The diagnosticity and overlap characteristics of the transfer items are also true with respect to the control study set. For example, the 3-1(3 overlap) transfer items 11313 and 11413 overlap the first control study item, 11112, on the first, second, and fourth features and with the second control study item, 11223, on the first, second, and fifth features. Thus, these two transfer items share three features with both items in their respective control study pair as they did with both items in their corresponding generalize study pair. The critical difference is that the 3-overlap with the generalize study pair matched a category generalization, whereas the 3-overlap with the control study pair did not. A comparison of this control study pair with its other transfer items will indicate that the same relations described above for the generalize study set hold for the control study set.

There were 16 items in each study set and in each transfer item type, half Club 1 members and half Club 2 members.

For blocked presentation, the 16 study items were divided into four groups of four items each. Each group consisted of one Club 1 study pair and one Club 2 study pair. The order of presentation within each group of four items was permuted and the final sequencing of the four groups for study was randomly determined. This method assured that, in the generalize condition, generalizable items were separated by at most two intervening items. Random presentation was realized as a pseudo-random ordering of the items using a method similar to the one described above for blocking. The difference was that the Club 1 items and the Club 2 items combined into a group of four were selected from different study pairs. The actual ordering of the items differed from trial to trial within the constraints of the blocking and random ordering algorithms.

Apparatus. The experiment was controlled by a PDP 11/34 computer. Subjects were seated in individual rooms, each of which contained a CRT screen on which the stimuli were displayed.

Procedure. The experiment was divided into two phases, a study phase and a test phase. For the study phase, subjects were told that their task was to learn to classify 16 people as either "Dolphin Club" or "Koala Club" members on the basis of their description (club names were chosen to

correspond to the terminal response keys "d" and "k"). To encourage subjects to attend to all five features, they were told that club membership was determined in a complex fashion and that there was no bias with respect to membership on the basis of a single feature. They were also encouraged not to formulate and test hypotheses during learning, but to concentrate on memorizing each description with its club assignment. The study items were presented in blocks of 16. One pass through all 16 items constituted one trial. The learning criterion was set at one correct pass through all the 16 items, i.e., one 100% accurate trial. Subjects kept cycling through the 16 items until they reached this criterion.

The study items were presented one at a time in the middle of the terminal screen. Subjects hit either "d" or "k" to classify a person as a Dolphin Club member or a Koala Club member, respectively. As soon as a response was entered, feedback of the form "Right, Dolphin (Koala) Club" or "Wrong, Dolphin (Koala) Club" appeared on the screen. The description, the subject's response, and the feedback remained on the screen for 10 seconds. The screen then erased and the next item was presented. A 10 second response-time limit was set. If the subject did not classify the item within 10 seconds, the correct club membership appeared, followed by the 10 second study time. Subjects were informed that failure to respond within 10 seconds counted as an error. At the end of each pass through the 16 items, subjects were told their accuracy for that trial. Rest breaks were spaced after every fourth trial.

After reaching the learning criterion, subjects began the test phase. They were told their task was to classify a new set of people as quickly as possible without sacrificing accuracy. Both the study items and the transfer items were presented during the test phase in a different random order for each subject. The test items were presented one at a time in the center of the screen and subjects hit either "d" or "k" to classify the description. After the subject classified the item, the word "confidence" appeared on the screen. Subjects assigned a confidence rating to their judgment, ranging from 1 ("not at all confident") to 5 ("absolutely confident"). Subjects were informed that the confidence rating was not timed and were encouraged to make sure it accurately reflected how confident they felt about their judgment. The description and the subject's response remained on the screen until the confidence rating was made. The screen then erased and the next item was presented. Accuracy and confidence ratings were recorded for each classification.

Results

The mean number of trials to criterion in the study phase was: 12.05 for generalize-blocked; 13.50 for generalize-random; 16.95 for control-blocked; and 15.85 for control-random. The effect of study set was significant [$F(1,76) = 6.3$, $p = .014$]. Newman-Keuls tests indicated that the generalize-blocked and the control-blocked conditions differed significantly, but the difference between the study set conditions with random presentation was not significant by this test.[1] Although learning in the generalize-blocked condition was faster than in the generalize-random condition in the predicted direction, neither the blocking manipulation nor its interaction with study set was significant [F's$(1,76) < 1.0$]. Since both study sets had equivalent ratios of diagnostic to nondiagnostic values on each feature, faster learning in the generalize conditions could not be attributed to the utilization of independent, diagnostic cues.

Confidence scores were computed as the mean of a subject's confidence ratings on correct classifications minus his confidence ratings on incorrect classifications for a given test item type. Thus, confidence scores range from -5 to +5. Accuracy on the study items at retest for generalize-blocked was 88%, for generalize-random, 85%, for control-blocked, 81%, and for control-random, 79%. The mean confidence ratings of study items in these four conditions was 3.16, 3.38, 2.73., and 2.84, respectively. While suggestive, the variation among these conditions on accuracy and confidence was not significant. The less than perfect performance on study items after reaching criterion during study probably reflects both successful guessing to reach study criterion and the subject's forgetting of his or her decision rules in the face of interfering transfer items.

Table 2 presents the mean accuracy and confidence rating for each transfer item type within each condition. Analyses of both the accuracy and confidence data for transfer items revealed a significant advantage for the generalize condition over the control condition [F's$(1,76) = 26.1$ and 24.5, respectively, $p < .001$]. For every transfer item type, subjects in the generalize condition were more accurate and more confident than subjects in the control conditions. There was a significant effect of type of transfer item on accuracy [$F(3,228) = 70.6$, $p < .001$] and confidence [$F(3,228) = 84.3$, $p < .001$]. Newman-Keuls tests on both the accuracy and confidence means revealed that all pairwise comparisons of item types differed significantly except the comparison of 2-2(2 overlap)

[1] All Newman-Keuls tests reported were significant at the .05 level.

Table 2

Experiment I:

Mean Accuracy and Confidence Scores on Transfer

Items as a Function of Study Material and Presentation Order

	Generalize		Control		
	Blocked	Random	Blocked	Random	Mean
Accuracy[a]					
3-1(3-overlap)	79	77	68	69	73
3-1(2-overlap)	70	71	56	64	65
2-2(3-overlap)	58	55	44	47	51
2-2(2-overlap)	54	57	44	43	50
Mean	65	65	53	56	60
Confidence					
3-1(3-overlap)	2.43	2.53	1.49	1.83	2.07
3-1(2-overlap)	1.66	1.77	0.57	1.37	1.34
2-2(3-overlap)	0.63	0.45	-0.47	-0.24	0.09
2-2(2-overlap)	0.51	0.56	-0.41	-0.57	0.03
Mean	1.31	1.33	0.30	0.60	.88

[a]percent correct

items and 2-2(3 overlap) items. The predicted study set by transfer item interaction did not occur. The blocking manipulation had no appreciable effect on either accuracy or confidence nor did it enter into any significant interactions.

Discussion

The results of Experiment I indicate that transfer to new category exemplars is facilitated when studied exemplars yield generalizations. In addition, the learning process itself is facilitated when generalizations can be formed between items being learned. The instance-only model cannot account for the effects of generalization on transfer performance. However, the generalization theory is not unequivocally supported, since we failed to find an interaction of study set with transfer item type. For both the generalize and control conditions, classification performance was a direct function of similarity to studied exemplars: The less similar transfer items were to studied items, the worse classification performance was. Under the view that a generalization must match a test item perfectly to apply, the generalization theory would predict good performance on the 3-1(3) transfer items, to which the category generalizations apply, and equally poor performance on all other transfer item types, to which the generalizations do not apply.

A careful post-experimental examination of the stimuli uncovered some unintended variation. Although the test items had satisfied our overlap constraints with the intended study pairs, they had a number of spurious overlaps with other study pairs. For example, while a 2-overlap item did in fact have only 2-features in common with each of its corresponding study items, it may have overlapped on three features with some other study items. To assess the extent of these spurious overlaps, we computed an overlap score for each test item to the generalize set and to the control set in the following manner. Each transfer item had two overlap scores. Its positive overlap score represented how similar it was with study items in its assigned category. Its negative overlap score represented how similar it was with study items in the alternative category. For each transfer item, we tabulated the frequency of 5-, 4-, 3-, 2-, and 1-overlaps it shared with all the study items in its assigned club (e.g., each Club 1 transfer was compared with all the Club 1 study items). Using the metric advocated by Medin and Schaffer (1978), these frequencies were weighted by the square of the amount of overlap they represented (e.g., the number of 3 overlaps were weighted by 9, the number of 2 overlaps by 4, and so on) and summed. This was the transfer item's positive overlap score, its similarity to

Table 3

Experiment I:

Mean Overlap Scores for Study Items and

Transfer Items as a Function of Study Condition

	Study Condition	
	Generalize	Control
Test Item		
Study[a]	22.6	23.5
3-1(3-overlap)	17.2	15.1
3-1(2-overlap)	10.1	2.7
2-2(3-overlap)	17.7	2.5
2-2(2-overlap)	0.5	-0.5

[a]This score is a measure of the amount of interitem similarity among the study items themselves.

study items in the category to which it was assigned. A transfer item's negative overlap score was computed in the same way, except that the transfer item was compared with study items in the alternative club (e.g., each Club 1 transfer was compared with all the Club 2 study items). A transfer item's final overlap score was the difference between its positive and negative overlap scores. Each transfer item had an overlap score for both the generalize study materials and the control study materials.

Table 3 gives the mean overlap score for each transfer item type with the generalize study set and with the control study set. The means for the study items represent their inter-item similarity. Note that, using this metric, the control study items had virtually the same amount, if not more, inter-item similarity than did the generalize items. Given these equivalent inter-item similarity scores for the two types of study materials, an instance-only model would be at a loss to explain the significantly faster learning of generalize study items.

To determine whether our results would stand with these spurious overlaps taken into account, we performed analyses of covariance using transfer items as the random factor and the individual transfer item overlap scores as the covariate. The same pattern of results emerged with these covariance analyses. Both study set and item type significantly affected accuracy [$F(1,59) = 6.3$, $p = .015$, and $F(3,59) = 21.7$, $p < .001$]. The interaction did not approach significance. The effect of these factors was significant with respect to confidence ratings as well [$F(1,59) = 7.5$, $p = .008$ and $F(3,59) = 20.3$, $p < .001$]. Consistent with the analyses of variance, generalize-condition subjects were more accurate and confident across item types than were control-condition subjects, and performance on transfer items decreased with decreasing similarity to studied items.

Experiment I did find an effect of generalize versus control study materials on subsequent transfer classification performance. The next experiment will focus on the effect of blocked versus random presentation and the third experiment will focus on the effect of generalize versus control study materials. By examining each of these factors one at a time, we were able to avoid the design constraints that led to the large amount of uncontrolled variation in overlap between study items and transfer items in Experiment I. In addition, by focusing on a single factor at a time, we were able to perform a more powerful manipulation of each variable and in addition, get the added statistical power of a within-subjects design.

Experiment II

In Experiment II, we contrasted two generalize conditions, one in which forming generalizations might be facilitated by blocking and one in which forming generalizations was hindered by random presentation of instances. To enhance the effect of blocking, we increased the ratio of items to generalizations, so that a given generalization accounted for three exemplars per category rather than just two. The strength of the resulting generalizations should be greater than in the previous experiment and the potential for blocking to have an effect should be greater. To increase the statistical power of the experiment, we made the presentation-order manipulation a within-subjects factor by running a two-phase experiment. In phase 1, subjects studied generalizable items presented either blocked or randomly. In phase 2, subjects studied a different set of generalizable materials in the alternative presentation order.

Method

Subjects. Forty-three members of the Carnegie-Mellon University community received Psychology course credit and/or $3.00 an hour for their participation in the two-hour experiment.

Materials and Design. Two sets of stimulus items were used to create 5-feature descriptions of people to be classified as Dolphin or Koala Club members and of "space invaders" to be classified as "friendly" or "hostile." Two sets were constructed to be used in the two phases of the experiment so that the form of the generalizations (i.e. which of the five features comprised the generalizations) would be different in each of the two phases (see Table 4 for stimulus set A and the Appendix for stimulus set B).

Each feature had six values. The club member features and values were: *job*: unemployed, self-employed, government, private-firm, military, retired; *religion*: Catholic, Jewish, Episcopalian, Baptist, Mormon, Lutheran; *hobby*: stamps, coins, painting, gardening, chess, reading; *musical taste*: classical, jazz, rock, disco, folk, country; *sport*: volleyball, basketball, bowling, squash, racquetball, handball. The space invader features and values were: *color*: purple, red, blue, green, yellow, brown; *skin*: metallic, furry, spiny, scaly, translucent, luminescent; *appendage*: claws, antennae, horns, wings, tentacles, tail; *home planet atmosphere*: radon, neon, helium, xenon, argon, krypton; *base of operations*--Venus, Mars, Jupiter, Saturn, Uranus, Pluto. As in Experiment I, the ordering of features in the description and the assignment of descriptive values to numeric values was randomly

Table 4

Experiment II:

Set A Generalize Study Items and Transfer Items

Study Items

Category 1	Category 2
11132	44423
11144	44411
11121	44434
14312	41243
24412	31143
44112	11443
32223	23332
43223	12332
21223	34332

Transfer Items

4-overlap		3-overlap		2-overlap[a]	
Category 1	Category 2	Category 1	Category 2	Category 1	Category 2
11111	44444	11115	44445	*12113	*43442
11114	44441	11116	44446	*13113	*42442
11134	44421			*14224	*41331
14212	41343	54212	51343	12214	43341
24112	31443	54612	51643	13211	42344
24212	31343			22114	33441
22223	33332	15223	45332	23114	32441
12223	43332	26223	36332	24141	31414
13223	42332			23214	32341

[a] Asterisks indicate items which partially match a category generalization

determined for each subject. There were nine items in each category. Study items were generated in sets of three. The three items in a set shared three features in common. Thus, there were three generalizations per category. The three category 1 generalizations were: *111--*, *-4-12*, and *--223*. The corresponding category 2 generalizations were *444--*, *-1-43*, and *--332*.

Three types of transfer items were constructed: 4-overlap, 3-overlap, and 2-overlap. The 4-overlap items shared the generalization yielded by one of the study set triplets plus a fourth feature with some of the items in the triplet. For example, the study item triplet 11144, 11121, and 11132 yields the generalization *111--*. The 4-overlap item 11134 shares the *111--* generalization with each of these three study items. It also overlaps the first study item on the fifth feature and the third item on the fourth feature. There were 18 (nine per category) 4-overlap transfer items.

The 2-overlap transfers overlapped on only two features with any study item in their respective category. A computer program generated all possible 2-overlaps items for each stimulus set. From this set, we selected the items which had a relatively high (four or more) number of positive 2-overlaps and a relatively low (two or fewer) number of negative 2-overlaps. Some of these 2-overlap transfers had the property that the two features they shared with a study item matched part of a category generalization. For example, the three study items 11144, 11121, and 11132 have the generalization *111--* and the 2-overlap transfer item 12113 overlaps on the first and third features of each of these items and with the generalization as well. Two overlap items which had this property were designated as 2-overlap partial matches (2(PM)-overlaps), since they matched two-thirds of a category generalization. In contrast, a 2-overlap item such as 12214 also overlapped three study items (11144, 14312, and 32223) on two features, but none of these 2-feature overlaps partially matched any of the category generalizations. For stimulus set A, 6 out of 18 2-overlap transfers were partial matches. For stimulus set B, 10 out of 18 2-overlap transfers were partial matches. The partial matches are starred in Table 4 and in the Appendix.

A third type of transfer item, 3-overlap transfers, was also included. These items matched one of the 3-feature category generalizations yielded by one study item triplet. However, they were qualitatively different from the other transfer items since one of their non-overlapping features had values which were not used in any of the study items. In other words, the study items used only four of the six possible values on a given feature, but the remaining two values were used to construct the

3-overlap transfer items. For example, the 3-overlap item 11115 overlaps each of the study items 11144, 11121, and 11132 on the *111*-- generalization, but the value 5 on the fifth feature was never used in any study item. It was necessary to use new values in order to construct items which overlapped on only the 3-feature generalization for one category while not overlapping on three features with an item in the alternative category. Since 3-overlap transfers contained never-studied values, they were always presented as the last items in the transfer test. This was done to insure that performance on 4- and 2-overlap transfers was uncontaminated by any "surprise" effects these items might generate. There were 12 (six per category) 3-overlap transfers in each stimulus set A and B. To summarize, category generalizations could be used to classify 4-overlap and 3-overlap transfer items; according to the full match view, they would not be helpful in classifying either 2-overlap or 2(PM)-overlap transfer items.

The presentation factor (blocked versus random) varied within subjects. In one phase, a subject's study items were blocked and in the other, the study items were presented randomly. For blocked presentation, two study item triplets, one from category 1 and one from category 2, were randomly selected to be combined as a group of six items, whose order was then permuted. A second pair of study item triplets was selected, combined, and permuted as a group of six. The final pair of triplets were then permuted. These 18 study items were then presented on one trial in this order. This method assured that the three items yielding a given generalization were clustered relatively close in the presentation sequence. For random presentation, items were also sorted into three groups of six, but the items in a given group of six came from each of the six different triplets. None of the three category 1 items in a group of six were from the same category 1 triplet, so they did not yield any category generalizations among themselves. Thus, there were no generalizable pairs in any block of six items. The order of the six items in each group was permuted and the 18 items were presented in this order. The actual ordering of items in the blocked and random conditions varied from trial to trial, within the constraints of their respective presentation algorithms.

Apparatus and Procedure. The apparatus and procedure were the same as described for Experiment II. The only difference was the learning criterion. To assure that subjects would complete both phases of the experiment in the allotted time, the learning criterion was set at two 85% correct passes through the 18 study items. If a subject did not reach this criterion after 14 passes, she or he

moved onto the transfer test of the phase.

Results

There were 14 cases of failing to reach the learning criterion. Six of these were due to three subjects who did not reach criterion in either phase of the experiment. Of the remaining eight cases, five occurred with blocked presentation and three occurred with random presentation. The mean number of trials to criterion for learning was 8.2 in the blocked phase and 9.5 trials in the random phase. The advantage of blocked presentation was significant [$F(1,41) = 4.4$, $p = .041$]. When the order of phases, blocked-random or random-blocked, is considered as a between-subjects factor, blocked-random subjects averaged 8.9 trials on their blocked (first) phase and 9.2 trials on their random (second) phase. Random-blocked subjects averaged 9.7 trials in their random (first) phase and 7.5 trials in their blocked (second) phase. The effect of phase order was not significant and did not interact with presentation: Subjects always learned faster in their blocked phases than in their random phases.

In the blocked condition, the mean accuracy on study items at retest was 79% and the mean confidence was 2.58. For random phases, these scores were 79% and 2.56. Table 5 presents the mean accuracy and confidence scores for transfer items.

Analyses of variance on the transfer data included the order of blocked and random phases as a between-subjects factor. There were 21 subjects whose phase order was blocked-random and 22 subjects whose phase order was random-blocked. The 2-overlap transfer items were partitioned into 2(PM)-overlaps and 2-overlaps, making a total of four transfer item types. There was a main effect of presentation on accuracy [$F(1,41) = 5.2$, $p = .029$]. Subjects' mean accuracy on transfer item was 73% in blocked phases but only 67% in random phases. The means in Table 5 show that accuracy varied greatly as a function of transfer item type, an effect which was highly significant [$F(3,123) = 31.5$, $p < .001$]. Not surprisingly, subjects were most accurate on 4-overlap transfers and least accurate on 2-overlap transfers. Newman-Keuls tests on the blocked-phase accuracy means indicated that accuracy on each of the two types of high (4 and 3) overlap transfers was significantly higher than accuracy on each of the two low (2 and 2(PM)) overlap transfers. Similarly, random-phase accuracy means for 4- and 3-overlap items were significantly higher than accuracy on all the 2-overlap items. It is interesting to note that subjects' accuracy on on 2(PM)-overlap items, which

Table 5

Experiment II:

Mean Accuracy and Confidence Scores on Transfer

Items as a Function of Presentation Order

	Presentation Order		
	Blocked	Random	Mean
Accuracy[a]			
Transfer Item			
4-overlap	83	78	81
3-overlap	78	72	75
2(PM)-overlap	69	64	67
2-overlap	61	53	57
Mean	73	67	70
Confidence			
Transfer Item			
4-overlap	2.80	2.42	2.61
3-overlap	1.86	1.63	1.75
2(PM)-overlap	1.24	1.03	1.14
2-overlap	0.87	0.36	0.62
Mean	1.69	1.36	1.53

[a]percent correct

partially matched two of the three features of a category generalization, was significantly higher than their accuracy on 2-overlap items. Apparently, having a partial overlap with the generalization led to an advantage.

Subjects were also more confident in blocked presentation conditions, but unlike the accuracy results, this effect did not reach statistical significance [$F(1,41) = 2.81$, $p = .1$]. There was a significant effect of transfer item type on confidence scores [$F(3,123) = 39.1$, $p < .001$]. Newman-Keuls tests revealed that all pair-wise comparisons of confidence means for the transfer item types were significant, with the exception of the 3-overlap and 2(PM)-overlap contrast.

Discussion

Support for the proposal that generalizations are formed during learning and utilized during transfer comes from a number of sources in Experiment II. First, learning was faster when generalizable items were blocked than when they were randomly ordered. Second, transfer performance was higher in the blocked phases than in the random phases. However, there was no interaction of presentation mode with test item type: The effect of decreasing similarity of test items to study items was the same in both blocked and random conditions. While blocking items may have facilitated forming generalizations, the effect of transfer item type in the random condition suggests that some category generalizations were formed even when generalizable items were randomly ordered. The third piece of (unexpected) evidence which argues for the existence of generalizations is the better transfer to 2(PM)-overlap items relative to 2-overlap items. Subjects were more accurate on 2-overlap transfers which partially matched a generalization than on those which did not, a result the generalization model would not have predicted.

To better evaluate this advantage for partial matches to generalizations, we used two metrics to determine whether 2(PM)-overlap items were qualitatively different from the 2-overlap items on some other dimension besides partially matching a category generalization. First, we classified 2(PM)- and 2-overlap items according to the number of diagnostic values in each item. For stimulus sets A and B combined, the frequency of 2(PM) overlap items with two, three, four, or five diagnostic values was 0, 4, 12, and 4, respectively. For 2-overlap items, it was 2, 12, 6, and 0, respectively. We can evaluate these frequency distributions by considering that each study item had exactly three diagnostic values: Twenty of the 2(PM)-overlap items and sixteen of the 2-overlap items had three or more diagnostic

values. On this measure, then, the 2(PM)-overlap transfers had a slight advantage. Second, we computed an overlap score for each of the 2(PM)-overlap and 2-overlap items, using the method described in Experiment I. Although each 2-overlap and 2(PM)-overlap transfer had only 2- and 1-overlaps with the study items, this similarity measure is based on the frequency with which the overlaps occur. The mean overlap score for the 2(PM)-overlap items, sets A and B combined, was 9.25. For the 2-overlap items, the mean overlap score was 8.0. While the differences between the 2(PM)-overlap items and the 2-overlap items on these two measures are small, they could account for the performance differences we found. We performed an analysis of covariance using item type (2(PM)-overlap versus 2-overlap) as the random factor and individual transfer item overlap scores as the covariate. With the analysis, the effect of item type on accuracy approached significance [$F(1,33) = 3.6$, $p = .06$]. This analysis suggests that the 2(PM)-overlap advantage did not occur simply because these items had *more* 2-feature matches with study instances than did the 2-overlap items, but because some of those 2-feature matches also partially matched a generalization. In other words, there seemed to be an effect for similarity to higher-order information. However, while the overall performance differences on low overlap items which do and do not partially match category generalizations are suggestive, the role of partial matches to generalizations in classification judgment warrants a more controlled investigation of its own.

Experiment III

Experiment III was designed as an attempt to replicate Experiment I's result that transfer to new items is better if studied items yielded category generalizations than if they did not. As in Experiment I, we contrasted a generalize study set, whose items yielded 3-feature generalizations, with a control set, whose items did not yield category generalizations. In Experiment I, there turned out to be some differences between the two sets of material in terms of similarity between study and transfer items. There still was an advantage for the generalize condition, covarying out this similarity, but it would be desirable to show a generalize advantage for having studied generalizable materials which were more equivalent to control materials in terms of inter-item similarity of study and transfer items. We discovered that we could not generate control and generalize study material which were equally similar to the transfer items and which satisfied overlap and cue validity constraints, if we used the same transfer item set for both generalize and control study sets as we had in Experiment I. Instead,

Table 6

Experiment III:

Study Items and Transfer Items

for Generalize and Control Conditions

Study Items

Generalize		Control	
Category 1	Category 2	Category 1	Category 2
11235	44325	11235	44325
11241	44314	12141	43414
12311	43244	11124	44431
12414	43141	21313	34242
23113	32442	31112	24443
24111	31444	42121	13434
31122	24433	24411	31144
41121	14434	13212	42343

Transfer Items

Generalize		Control	
Category 1	Category 2	Category 1	Category 2
11213	44342	11241	44314
11224	44331	12135	43425
12213	43342	21123	34432
12115	43445	11314	44241
22115	33445	41111	14444
21114	34441	32122	23433
21124	34431	23211	32344
11123	44432	14412	41143

we designed a generalize study set and a control study set each with its own transfer item set and endeavored to make the relation between the transfer set and the study set as equivalent as possible for both the generalize and the control materials.

Method

Subjects. Forty members of the Carnegie-Mellon University community received Psychology course credit and/or $3.00 an hour for participation in the two-hour experiment.

Materials and Design. The club member and space invader materials described for Experiment II were used in conjunction with the generalize and control items presented in Table 6.

For both the generalize and control study items, there were four possible values for each of the first four features and five possible values for the fifth feature. Category 2 was constructed from category 1 by interchanging 1's with 4's and 2's with 3's. The value 1 on any feature was quite diagnostic of category 1 and the value 4 on any feature was quite diagnostic of category 2. Except for the fifth feature, the value 2 was somewhat diagnostic of category 1 and the value 3, of category 2. The values 2, 3, and 5 on the fifth feature were not diagnostic of either category.

There were eight study items per category. For the *generalize condition*, pairs of study items were constructed to overlap on three features. The four category 1 generalizations were: *112--*, *12-1-*, *2-11-*, and *-112-*. For the control condition, study items were also constructed in a pairwise manner. The items in a control study-pair shared only one feature in common. Using the metric described in Experiment I, the mean overlap score for the generalize study items, a measure of their inter-item similarity, was 28.8; for control study items, the mean overlap score was 25.1.

Only one type of transfer item, 3-overlap, was used; there were eight 3-overlap transfer items per category. One pair of transfer items was constructed for each pair of study items. For the generalize study set, each of the transfer items in a pair overlapped each of the items in its corresponding study pair on three features. These three features were the generalization yielded by the study pair. For example, the study pair 11235 and 11241 overlap on the first, second, and third features, yielding the generalization *112--*. The two transfer items 11213 and 11224 overlap each of the study items on the first three features and are classifiable by the *112--* generalization. A transfer item pair overlapped on three features only with the items in its corresponding study pair, i.e., there were no spurious 3-feature overlaps between a transfer item and a third study item.

Pairs of transfer items were constructed in the same way for the control study set. Each transfer item in a pair overlapped with two study items on three features in its corresponding study pair, but a different three features for each item. For the control study pair, 11235 and 12141, the transfer item 11241 overlaps the first study item on the first, second, and third feature and with the second study item on the first, fourth, and fifth features. The second transfer item 12135 overlaps the first study item in the first, fourth, and fifth feature and on the second item on the first, second and third item. As in the generalize study set, a control transfer pair shared three features only with its corresponding study pair.

An overlap score was computed for each transfer item by tabulating the frequency of 3-, 2-, and 1-overlaps with study items in its assigned category (positive overlap) and with study items in the alternative category (negative overlap). As in Experiment I, we multiplied the frequency of each overlap type by the square of the overlap, and summed the results. The positive overlap score minus the negative overlap score gave the final overlap measure. The mean overlap score for the generalize transfer items was 23.6 and for the control transfer items, it was 18.6. Since this difference was larger than we had hoped to achieve, transfer item overlap scores were subsequently used as the covariate in analyses of covariance on the recall data.

Study materials, generalize or control, was varied within subjects. The experiment was run in two phases. The order of phases, generalize-control or control-generalize, and the assignment of club member or space invader descriptions to generalize and control materials, was counterbalanced across subjects.

The generalization model predicts better transfer performance in the generalize condition, in which transfer items are classifiable by category generalizations, than in the control condition. An instance-only model would predict no difference between generalizable and control conditions on transfer performance.

Apparatus and Procedure. The procedure was identical to that of Experiment II. Both generalize and control study items were presented in a blocked order, using the method described in Experiment II. After each transfer test, subjects filled in brief questionnaires in which they described (a) what strategies they used to learn the study items and (b) their impressions of what determined category membership.

Results and Discussion

There were 26 cases of failure to reach learning criterion. Five subjects did not reach criterion for either their generalize or control phase. Of the remaining 16 cases, four occurred in the generalize phase and twelve occurred in the control phase. Subjects took 9.55 trials to reach learning criterion with generalize materials and 10.68 trials with control materials. This effect approached significance [$F(1,38) = 3.65$, $p = .06$]. Learning was faster in the second phase, regardless of materials, as revealed by a significant study material by phase order interaction [$F(1,38) = 13.01$, $p = .001$]. Generalize-control subjects took 10.3 trials in their first (generalize) phase and 9.3 trials in their second (control) phase. Control-generalize subjects took 12.1 trials in their first (control) phase and 8.8 trials in their second (generalize) phase. The speed-up across phases for subjects going from generalize to control materials was 1.0 trials, but for subjects going from control to generalize materials, the decrease in learning time was more than three trials. These learning phase data replicate the findings of Experiment I: Learning was facilitated when the study items afforded category generalizations, even when the two sets of study materials were approximately equivalent in terms of inter-item similarity.

For generalize materials, the mean accuracy and confidence on study items at retest was 82% and 2.87, respectively. For control materials, these scores were 82% and 2.91.

Table 7

Experiment III:

Mean Accuracy and Confidence Scores on Transfer

Items as a Function of Study Set and Phase Order

	Phase Order		
	Generalize-Control	Control-Generalize	Mean
Accuracy[a]			
generalize	74	82	78
control	72	70	71
Confidence			
generalize	2.11	2.90	2.50
control	1.95	1.79	1.85

[a]percent correct

Table 7 presents the mean accuracy and confidence scores as a function of phase-order and materials for the transfer items. There was a main effect of item type (study versus transfer) on accuracy [$F(1,38) = 25.1$, $p < .001$]. Not surprisingly, subjects were less accurate on new transfer items than on studied items, indicating some effect of memory for specific instances. While subjects were equally accurate on generalize and control study items, they differed significantly in accuracy on the 3-overlap transfer items as a function of study material ($p < .05$). With generalize materials, subjects were 78% accurate on transfer items, while with control materials, they were 71% accurate on transfer items. This interaction of study materials with test item type was significant [$F(1,38) = 4.6$, $p = .038$]. Phase order did not significantly affect accuracy, but the phase order by study materials interaction approached significance [$F(1,38) = 3.5$, $p < .06$]. This reflected the trend that control-generalize subjects were 12% more accurate in their second (generalize) transfer test than they were in their first (control) transfer test, whereas generalize-control subjects were only 2% more accurate in their second (control) phase than in their first (generalize) phase. At the very least, these data suggest that the benefit of practice with the task is contingent on the nature of the materials.

Similar effects emerged for confidence ratings. The item effect was significant [$F(1,38) = 30.3$, $p = .001$]. The mean confidence score on transfer items for generalize materials, 2.50, was significantly higher than the mean confidence score on these items given control materials, 1.85 ($p < .05$). The interaction of study materials with item type was significant [$F(1,38) = 8.8$, $p = .005$]. Study material interacted with phase order [$F(1,38) = 6.2$, $p = .017$]. The mean increase in confidence on the second phase relative to the first was 1.11 for control-generalize phase order, but only .16 for the generalize-control phase order.

As in Experiment I, we performed an analysis of covariance using items as the random factor with the overlap scores serving as the covariate. Under this analysis, the adjusted accuracy and confidence means for the transfer items in the generalize and control phases still differed significantly ($p < .05$).

When we examined the subjects' post-experimental reports, we found little evidence for awareness of generalizations or parts of generalizations. There was certainly no case in which a subject reported all six 3-feature generalizations which occurred in his or her study items. When asked what determined category membership, most subjects listed single features. A few subjects showed

sensitivity to configurations of features or contingency relationships (e.g., "Dolphin members were Lutheran and collected stamps, unless they liked jazz. . . then they were Koalas"). We checked each subject's report for the generalize phase to see how well their rules matched the six generalizations which actually appeared in his or her study items. Out of the 30 subjects for which we had protocols, one subject reported two complete generalizations; another subject mentioned one. There were seven subjects reporting two-thirds of some of the generalizations. However, these subjects, like the majority, also reported feature combinations which did not correspond at all to the generalizations. In general, subjects were either unaware of the category generalizations or unable to articulate them.

General Discussion

Perhaps the best testimony to the success of these experiments is that none of the theories we reviewed in the introduction has emerged unscathed. Experiments I and III provided ample evidence for the advantage of generalizations during initial learning and in transfer. While Experiment II did not directly contrast a control versus generalize condition, the contrast between blocked and random presentation would only have impact if subjects were forming generalizations. This means that subject performance is not just a direct function of similarity between study instances and test instances. Rather, inter-item similarity among the study items is important in creating generalizations that can be used to categorize transfer items. These results, as well as the advantage in both speed of learning and transfer of the generalize condition over the control condition rule out the pure instance-only theories (Medin & Schaffer, 1978).

The prototype models do suppose that central tendencies are extracted and used to categorize test instances. However, they assume a single central tendency which implies that distance from the central tendency should be the only relevant variable. Numerous experiments have already disconfirmed this prediction (e.g., Hayes-Roth & Hayes-Roth, 1977; Medin and Schaffer, 1978) as commented upon in the introduction. Experiment I showed that, for our particular paradigm too, there is an effect of degree of overlap with individual study instances, holding number of diagnostic features (i.e., central tendency) constant.

Our results also rule out most of the feature-set models (Reitman & Bower, 1973; Neumann, 1974; Hayes-Roth and Hayes-Roth, 1977) in that they have no role for a generalization process. Both Neumann's (1974) model and Reitman and Bower's (1973) model were designed to predict

recognition ratings; neither makes classification predictions. Hayes-Roth and Hayes-Roth's (1977) property set model, which is most similar to the generalization model we tested, does not predict the differences we found between generalize and control conditions. Their model predicts classification on the basis of a transfer item's most diagnostic property set. According to their model, an item's property sets are the powerset of all its values; each of our 5-feature items had 31 property sets. To make property set model predictions for our Experiment III task, we computed the following for each property set for each transfer item: (a) tabulated its frequency of occurrence among category 1 exemplars and among category 2 exemplars, (b) formed two ratios: the frequency of category 1 occurrences over all occurrences and the frequency of category 2 occurrences over all occurrences, and (c) found the largest ratio of all those computed--this signified the most diagnostic property set. If it is a category 1 ratio, the model predicts a category 1 classification for the transfer item. If it is a category 2 ratio, the model predicts a category 2 classification. If there is a tie for the largest ratio across categories, classification is not predicted.

In Experiment III, there are 12 generalize transfer items with a most-diagnostic property set for the correct category. There are 10 such control items. Performance is 81% accurate on these generalize items and 73% on these control items. Thus, when we consider only those transfer items for which the property set model makes a classification prediction, it still cannot account for the difference found between generalize and control conditions.

So this leaves only the ACT model. However, that theory as set forth in Anderson et al. (1979) also has numerous problems. A major difficulty is that it does not successfully classify a transfer item unless it is perfectly matched by some study item or by a generalization formed from study items. This leaves the theory at a loss to explain many results in the present experiments, such as how subjects could successfully categorize transfer items at all in the control condition where there were no generalizations or how they could at all categorize test items in the generalize condition that only partially overlapped with the generalizations. This produced the failure to obtain an interaction between treatment and transfer item types in Experiments I and II.

It seems that the major inadequacy with the ACT theory as formulated by Anderson et al. (1979) is its failure to allow items to be classified on the basis of partial matches. In response to these results and other considerations, the ACT pattern matcher has been augmented to permit partial matches to

both specific instances and to category generalizations. This means that there are two bases for similarity to lead to transfer in these categorization tasks. First, inter-item similarity can lead to category generalizations at study. In addition, similarity between transfer items and either study items or higher-order category information (generalizations) can serve as a direct basis for categorization. In fact, the current ACT model uses the same partial matching techniques for detecting similarities between study items to form generalizations as it uses for categorizing new items.

It is interesting here to consider the distinction proposed by Brooks (1977) between two modes of category learning: analytic and nonanalytic. In simplified terms, a learner is said to be in analytic mode if he or she is trying to form, test, and abstract rules about category membership. In nonanalytic mode, the learner attends to individual items, and subsequent classification of new items is accomplished via analogy to remembered items. This analogy approach is similar to the Medin and Schaffer's (1978) theory, but Brooks specifies the conditions under which either mode is likely to occur as well as their respective behavioral consequences. Since we believe we have found evidence for both (a) abstraction of higher-order category information and (b) similarity effects to both specific instances and this higher-order category information, it is interesting to evaluate whether our subjects were analytic or nonanalytic learners when given generalize materials. Using the criteria Brooks offers, we believe that, at least for Experiments I and III, our subjects' approach was nonanalytic. First, the instructions emphasized the complexity of the to-be-learned categories and that no single rule determined membership. The experimenter offered further admonitions against the testing of specific hypotheses on successive passes through the study items, warning that this would be a frustrating and unsuccessful endeavor. Second, memory for study items retested during the transfer tasks was quite high. Brooks notes that one consequence of analytic processing is poor memory for study items, since attention is devoted to uncovering consistencies among items rather than to learning individual items. Allowing for some interference effects, our subjects classified studied items fairly accurately, and the difference between old and new items was also reflected in higher confidence on studied items. Third, Experiment III subjects' ability to specify rules in their post-experimental reports was similar and poor across generalize and control conditions. For the generalize phase, the majority of their rules bore little resemblance to generalizations actually used.

These remarks suggest that our subjects approached and executed the task in a fairly nonanalytic

manner, regardless of the type of material they studied. Yet we found evidence for the use of category generalizations during transfer, not just analogy to studied items. The present data suggest that it may be both unnecessary and inappropriate for a theory of schema abstraction to choose between rule abstraction mechanisms and analogy mechanisms. Both specific instance information and higher-order category information may be available after nonanalytic processing of category exemplars and an analogy, partial-matching, or some such similarity-detecting mechanism may operate in the same way on both types of information.

Appendix

Experiment II: Set B Generalize Study Items and Transfer Items

Study Items

Category 1	Category 2
11321	44234
11441	44114
11211	44344
43121	12434
44122	11433
41124	14431
22233	33322
32234	23321
12232	43323

Transfer Items

4-overlap		3-overlap		2-overlap[a]	
Category 1	Category 2	Category 1	Category 2	Category 1	Category 2
11111	44444	11151	44454	*13112	*42443
11141	44414	11161	44464	*21112	*34443
11341	44214			*21412	*34143
42121	13434	42125	13435	*22112	*33443
41122	14433	46122	16433	*24212	*31343
42122	13433			22442	33113
22232	33323	52231	53324	13243	42312
22231	33324	62231	63324	21143	34412
32231	23324			23132	32423

[a] Asterisks indicate items which partially match a category generalization

References

Anderson, J. R. *Language, Memory, and Thought.* Hillsdale, N.J.: Erlbaum Associates, 1976.

Anderson, J. R., Kline, P. J., & Beasley, C. M. A general learning theory and its application to schema abstraction. ONR Technical Report (78-2), Yale University, 1978.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 331-350.

Brooks, L. Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1978.

Franks, J. J., & Bransford, J. D. Abstraction of visual patterns. *Journal of Experimental Psychology*, 1971, *90*, 6574.

Hayes-Roth, B., & Hayes-Roth, F. Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 321-338.

Medin, D. L., & Schaffer, M. M. Context theory of classification learning. *Psychological Review*, 1978, *85*, 207-238.

Neumann, P. G. An attribute frequency model for the abstraction of prototypes. *Memory & Cognition*, 1974, *2*, 241-248.

Posner, M. J., & Keele, S. W. On the genesis of abstract ideas. *Journal of Experimental Psychology*, 1968, *77*, 353-363.

Posner, M. J., & Keele, S. W. Retention of abstract ideas. *Journal of Experimental Psychology*, 1970, *83*, 304-308.

Reitman, J. S., & Bower, G. H. Storage and later recognition of exemplars of concepts. *Cognitive Psychology*, 1973, *4*, 194-206.

Rosch, E., & Mervis, C. B. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 1975, 7, 573-605.

Navy

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Robert Breaux
Code N-711
NAVTRAEQUIPCEN
Orlando, FL 32813

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Henry M. Halff
Department of Psychology,C-009
University of California at San Diego
La Jolla, CA 92093

1 LT Steven D. Harris, MSC, USN
Code 6021
Naval Air Development Center
Warminster, Pennsylvania 18974

1 CDR Charles W. Hutchins
Naval Air Systems Command Hq
AIR-340F
Navy Department
Washington, DC 20361

1 CDR Robert S. Kennedy
Head, Human Performance Sciences
Naval Aerospace Medical Research Lab
Box 29407
New Orleans, LA 70189

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

Navy

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OP01T
Washington DC 20370

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr William Montague
Navy Personnel R&D Center
San Diego, CA 92152

1 Library
Naval Health Research Center
P. O. Box 85122
San Diego, CA 92138

1 Naval Medical R&D Command
Code 44
National Naval Medical Center
Bethesda, MD 20014

1 Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

Navy

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Special Asst. for Education and
Training (OP-01E)
Rm. 2705 Arlington Annex
Washington, DC 20370

1 Office of the Chief of Naval Operations
Research Development & Studies Branch
(OP-115)
Washington, DC 20350

1 Captain Donald F. Parker, USN
Commanding Officer
Navy Personnel R&D Center
San Diego, CA 92152

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

Navy

1 Dr. Gary Poock
Operations Research Department
Code 55PK
Naval Postgraduate School
Monterey, CA 93940

1 Roger W. Remington, Ph.D
Code L52
NAMRL
Pensacola, FL 32508

1 Dr. Bernard Rimland (03B)
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Roger Weissinger-Baylon
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

Navy

1 Mr John H. Wolfe
Code P310
U. S. Navy Personnel Research and Development Center
San Diego, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 HQ USAREUE & 7th Army
ODCSOPS
USAAREUE Director of GED
APO New York 09403

1 DR. RALPH DUSEK
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 DR. FRANK J. HARRIS
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Col Frank Hart
Army Research Institute for the Behavioral & Social Sciences
5001 Eisenhower Blvd.
Alexandria, VA 22333

1 Dr. Michael Kaplan
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Army

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Dr. Marty Rockway
Technical Director
AFHRL(OT)
Williams AFB, AZ 58224

2 3700 TCHTW/TTGH Stop 32
Sheppard AFB, TX 76311

1 Jack A. Thorp, Maj., USAF
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

Other DoD

12 Defense Technical Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC

1 Dr. Craig I. Fields
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

1 Dr. Dexter Fletcher
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD.
ARLINGTON, VA 22209

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

Civil Govt

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Joseph I. Lipson
SEDR W-638
National Science Foundation
Washington, DC 20550

1 William J. McLaurin
Rm. 301, Internal Revenue Service
2221 Jefferson Davis Highway
Arlington, VA 22202

1 Dr. Arthur Melmed
National Intitute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

1 Dr. Frank Withrow
U. S. Office of Education
400 Maryland Ave. SW
Washington, DC 20202

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Anderson, Thomas H., Ph.D.
Center for the Study of Reading
174 Children's Research Center
51 Gerty Drive
Champiagn, IL 61820

1 Dr. John Annett
Department of Psychology
University of Warwick
Coventry CV4 7AL
ENGLAND

1 DR. MICHAEL ATWOOD
SCIENCE APPLICATIONS INSTITUTE
40 DENVER TECH. CENTER WEST
7935 E. PRENTICE AVENUE
ENGLEWOOD, CO 80110

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

1 Dr. Patricia Baggett
Department of Psychology
University of Denver
University Park
Denver, CO 80208

1 Mr Avron Barr
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

Non Govt

1 Dr. Lyle Bourne
Department of Psychology
University of Colorado
Boulder, CO 80309

1 Dr. John S. Brown
XEROX Palo Alto Research Center
3333 Coyote Road
Palo Alto, CA 94304

1 Dr. Bruce Buchanan
Department of Computer Science
Stanford University
Stanford, CA 94305

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. Pat Carpenter
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. William Chase
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Dr. Micheline Chi
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Non Govt

1 Dr. William Clancey
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Allan M. Collins
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, Ma 02138

1 Dr. Lynn A. Cooper
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

1 Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, DC 20036

1 Dr. Hubert Dreyfus
Department of Philosophy
University of California
Berkely, CA 94720

1 LCOL J. C. Eggenberger
DIRECTORATE OF PERSONNEL APPLIED RESEARC
NATIONAL DEFENCE HQ
101 COLONEL BY DRIVE
OTTAWA, CANADA K1A OK2

1 Dr. Ed Feigenbaum
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Mr. Wallace Feurzeig
Bolt Beranek & Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

Non Govt

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 Dr. Alinda Friedman
Department of Psychology
University of Alberta
Edmonton, Alberta
CANADA T6G 2E9

1 Dr. R. Edward Geiselman
Department of Psychology
University of California
Los Angeles, CA 90024

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Marvin D. Glock
217 Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Daniel Gopher
Industrial & Management Engineering
Technion-Israel Institute of Technology
Haifa
ISRAEL

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Harold Hawkins
Department of Psychology
University of Oregon
Eugene OR 97403

Non Govt

1 Dr. Barbara Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

1 Dr. Frederick Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

1 Dr. James R. Hoffman
Department of Psychology
University of Delaware
Newark, DE 19711

1 Glenda Greenwald, Ed.
"Human Intelligence Newsletter"
P. O. Box 1163
Birmingham, MI 48012

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Steven W. Keele
Dept. of Psychology
University of Oregon
Eugene, OR 97403

1 Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80302

1 Dr. David Kieras
Department of Psychology
University of Arizona
Tuscon, AZ 85721

1 Dr. Kenneth A. Klivington
Program Officer
Alfred P. Sloan Foundation
630 Fifth Avenue
New York, NY 10111

Non Govt

1 Dr. Stephen Kosslyn
Harvard University
Department of Psychology
33 Kirkland Street
Cambridge, MA 02138

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 Dr. Jill Larkin
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

1 Dr. Robert A. Levit
Director, Behavioral Sciences
The BDM Corporation
7915 Jones Branch Drive
McClean, VA 22101

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Erik McWilliams
Science Education Dev. and Research
National Science Foundation
Washington, DC 20550

1 Dr. Mark Miller
Computer Science Laboratory
Texas Instruments, Inc.
Mail Station 371, P.O. Box 225936
Dallas, TX 75265

Non Govt

1 Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Ave., Fourth Floor
Redondo Beach, CA 90277

1 Dr. Donald A Norman
Dept. of Psychology C-009
Univ. of California, San Diego
La Jolla, CA 92093

1 Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

1 Dr. Seymour A. Papert
Massachusetts Institute of Technology
Artificial Intelligence Lab
545 Technology Square
Cambridge, MA 02139

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. PETER POLSON
DEPT. OF PSYCHOLOGY
UNIVERSITY OF COLORADO
BOULDER, CO 80309

1 Dr. Steven E. Poltrock
Department of Psychology
University of Denver
Denver,CO 80208

1 MINRAT M. L. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

Non Govt

1 Dr. Fred Reif
SESAME
c/o Physics Department
University of California
Berkely, CA 94720

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 DR. WALTER SCHNEIDER
DEPT. OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61820

1 Dr. Alan Schoenfeld
Department of Mathematics
Hamilton College
Clinton, NY 13323

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Committee on Cognitive Research
% Dr. Lonnie R. Sherrod
Social Science Research Council
605 Third Avenue
New York, NY 10016

1 Robert S. Siegler
Associate Professor
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

Non Govt

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 DR. ALBERT STEVENS
BOLT BERANEK & NEWMAN, INC.
50 MOULTON STREET
CAMBRIDGE, MA 02138

1 Dr. David Stone
ED 236
SUNY, Albany
Albany, NY 12222

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. John Thomas
IBM Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

1 DR. PERRY THORNDYKE
THE RAND CORPORATION
1700 MAIN STREET
SANTA MONICA, CA 90406

1 Dr. Douglas Towne
Univ. of So. California
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

Non Govt

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Benton J. Underwood
Dept. of Psychology
Northwestern University
Evanston, IL 60201

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 DR. GERSHON WELTMAN
PERCEPTRONICS INC.
6271 VARIEL AVE.
WOODLAND HILLS, CA 91367

1 Dr. Keith T. Wescourt
Information Sciences Dept.
The Rand Corporation
1700 Main St.